

Figure 1

Nucleotide sequence of the *Prunus amygdalus HNL5* gene obtained by PCR amplification

The start codon (ATG) and stop codon of the open reading frame are printed in bold type, and the nucleotides in the intron regions are indicated in lower case letters. The peripheral sequences which have been attached via the PCR primers and which are not part of the *HNL5* gene are underlined. The splice sites of the introns were identified with the aid of the consensus sequence "GT...AG".

1 GGAAATTCACA ATATGGAGAA ATCAACAATG TCAGTTATAC TATTTGTGTT
51 GCATCTTCTT GTTCTTCATC TTCAGTATTC AGAGGTTTAC TCGCTTGCCA
101 ATACTTCTGC TCATGgtaaa ttccatctt cagtatcat ttaacagcaa
151 aetgtgtaga ttataatla agaaaactga cacaagtagt gcaagaasaa
201 agctaaftta gatgcaggtt gaasaaaatc ttcatctct tcacatatat
251 ttgcagATT TTAGCTACTT GAAGTTTGTG TACAACGCCA CTGATACAAG
301 CTCGGAAGGA TCATATGACT ACATTGTAAT CGGTGGAGGA ACATCAGGGT
351 GTCCATTGGC AGCAACTTTA TCAGAAAAAT ACAAGGTGCT TCTTCTAGAA
401 AGAGGCACTA TTGCTACAGA ATACCCGAAC ACGTTGACTG CAGATGGGTT
451 TGCATATAAT CTGCAGCAAC AAGATGATGG AAAGACGCCA GTTGAAAGST
501 TCGTGTCCGA AGATGGCATT GATAATGTGC GAGCCAGGAT CCTCGGTGGC
551 ACGACCATAA TCAATGCAGG CGTCTACGCC AGAGCTAACA TTTTATTCTA
601 TAGTCAAACA GGAATTGAAT GGGACCTGGA TTTGGTCAAT AAGACATATG

1046232.011602

651 AGTGGGTTGA AGACGCCATT GTGGTCAAGC CAAATAATCA ATCTTGGCAA
701 TCTGTTATAG GAGAGGGATT CTTGGAGGCG GGTATTCTTC CAGACAATGG
751 ATTTAGTTTG GATCACGAAG CAGGAAC TAG ACTCACC GGC TCAACTTTG
801 ACAATAATGG AACGCGACAT GCGGCTGATG AACTGCTTAA TAAAGGAGAC
851 CCTAATAACT TGCTAGTTGC AGTTCAGGCC TCAGTAGAGA AGATCCTCTT
901 CTCTTCCAAT ACATCAAgta tgttgcata gtgatatta atgtagctc
951 ctggttgc atgcgcact cgaattat tatattatc ttttaata
1001 ctaacagaat agtgaagt ctcatattc cctccatat ttccaaatt
1051 tccataaaca aaactccca atctccctc gtttagttg acaataata
1101 taagctattc tcaatgcag ATTTGTCAGC TATTGGAGTC ATATATACGG
1151 ATTCTGATGG AAACCTCAT CAGGCATTG TACGCGGTAA CGGAGAAGTT
1201 ATTGTTAGTG CAGGGACAAT CGGAACGCCT CAGCTTCTAC TACTTAGTGG
1251 CGTTGGACCA GAGTCTTACC TATCTTCTCT CAACATCACA GTTGTTCCAG
1301 CGAATCCTTA TGTTGGGCAG TTTGTGTATG ACAATCCTCG TAATTCATT
1351 AATATTTTGC CCCCAAATCC AATTGAAGCC TCTGTTGTAA CTGTTTATGG
1401 CATTAGAAGT GATTATTATC AAGTTTCTCT GTCAAGCTTG CCATTTTCCA
1451 CTCCACCCTT TAGTCTTTTT CGTACAACAT CTTACCCCTC CCCAAATTGG
1501 ACTTTTGCTC ATATTGTTAG CGAAGTTCCA GGACCATTGT CTCATGGTTC
1551 TGTCACGCTA AATTCATCAT CTGACGTGAG AATCGCTCCA AATATTAAAT
1601 TCAATTACTA TTCAAATTCC ACAGACCTTG CTAATTGTGT TAGCGGCATG
1651 AAGAAGCTTG GTGACTTATT AAGGACAAAG GCATTAGAAC CATATAAAGC

10046222.013607

1701 TCGAGATGTG CTGGGAATTG ACGGTTTCAA TTATTTGGGA GTACCTTTGC
1751 CAGAGAACCA AACAGATGAT GCATCCTTCG AAACATTTTG TCTAGATAAT
1801 GTAGCTTCAT ACTGGCATTa CCACGGTGGa AGCCTTGTTG GGAAAGTGCT
1851 TGATGACAGT TTCCGTGTTa TGGGGATCAa AGCATTACGC GTTGTTGATG
1901 CCTCCACTTT CCCTTACGAA CCAACAGCC ATCCTCAGGG CTTCTATCTG
1951 ATGTTAGGAA Ggtatgtget gcacacttc aeccactaga gattctcaat
2001 atttlttgtl tglgtaatg aactctctgc cgcattgcic tttttatta
2051 atocttaaaa ttttltgttl tggcgagGTA TGTGGGCCTT CAAATCCTGC
2101 AAGAAAGGTC AATCCGGTTG GAGGCTATTc ATAATATTCA AGAGTCCATG
2151 TGAAGAAATTC CG

1046232-01602

Figure 2

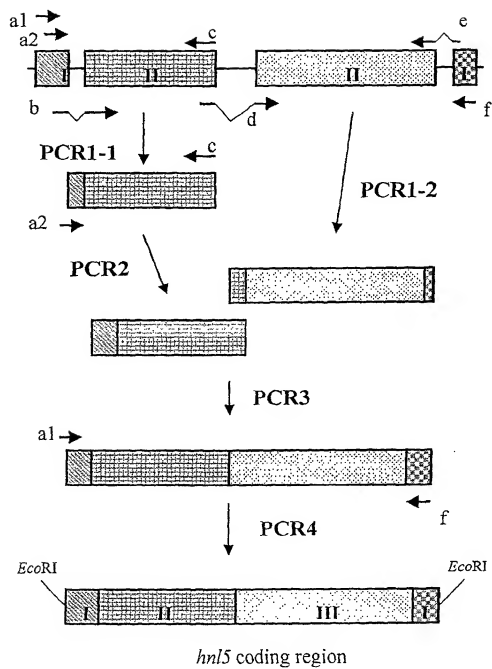


Figure 3:

Amino acid sequence of the *Prunus amygdalus* hydroxynitrile lyase (HNL5), derived from the nucleotide sequence of the *HNL5* gene. The signal sequence determined from sequence analysis is printed in bold type and the postulated processing site is indicated by an arrow. Possible glycosylation sites (PROSITE patterns) are underlined.

↓

MEKSTMSVILFVLHLLVLHLQYSEVHSLANTSAHDFSYLK FVY **NAT**DTSSSEGSYDI
VIGGGTSGCPLAATLSEKYKVLLERGTATEYPNTLTADGFAYNLQQQDDGKTPVE
RFVSEGDIDNVRARILGGTTIINAGVYARANISEFYSQTGIEWDLDLV **NK**TYEWVEDAI
VVKPN**NQSW**QSVIGEGFLEAGILPDNGFSLDHEAGRTLGTSTFDNN**NGTR**HAADELL
NKGDPNNLLVAVQASVEKILFSS**NT**SNLSAIGVIYTDSDGNSHQAFVRGNGEVIVSA
GTIGTPQLLLLSGVGFPESYLS**SLNT**TVVQPNPYVGQFVYDNPRNFINILPPNPIEASVV
TVLGIRSDYYQVLS**SSL**PFSTPPFSLFPTTSYPL**PNST**FAHIVSQVPGPLSHG**SVTLN**
SSSDVR!APNIKFNYY**SNTD**LANCVSGMKKLGDLRTKALEPYKARDVLGIDGFNY
LGVPLP**NOTD**DASFETFCLDNVASYWHYHGGSLVGKVLDDSRFVMGKALRVVD
ASTFPYEPNSHPQGFYLM**LG**RYVGLQILQERSIRLEAIHNQESM

20040323.011002

Figure 4:

Nucleic acid sequence of the DNA fragment coding for a secretory hybrid protein (PamHNL5xGOX) with HNL activity, consisting of sequences of the *Prunus amygdalus* HNL5 gene and the *Aspergillus niger* glucose oxidase gene.

gaattcatcatgcagactctcttgtgagctcgcttgggtccctcgctgcggccctgccac-
cactacatcaggagcaatggcattgaagcctacaacgccaatgatcaagctcggaaggatca-
tatgactacattgtaatcggtggaggaaacatcagggtgtccattggcagcaacttatcagaa-
aatacaagggtgctctcttagaaggaggcactatgtctacagaataccgaacacgtt-
gactgcagatgggtttgcatataatctgcagcaacaagatgatggaagacgccagttga-
aaggttcggtccgaagatggcattgataatgtgcgagccaggatcctcggtggcagcacca-
taatacatgcaggcgctctacgccagagctaacatttcatctatagtaacaacaggaatt-
gaatgggacctggatttggcacaataagacatagtggttgaagacgcatgtggt-
caagccaaataatcaatcttggcaatctgttataggagagggatcttggaggcggg-
tattctccagacaatggatttagtttggatcagcaagcaggaaatagactcacggct-
caactttgacaataatggaacgcgacatgcggctgatgaactgcttaataaggagagcc-
taataacttgcagttgcagttcaggccctcagtagagaagatcctctcttccaatacat-
caaatgttcagctattggagtcataatacggattctgatggaaactctcatcaggcattg-
tacgcggtaacggagaaagtattgttagtgagggaacaatcggaaacgcctcagctctac-
tcttagtggcggttggaaccagagcttacctatctctcaacatcacagttgttcagcc-
gaatccctatgttgggcagttgtgtatgacaatccctgtaatttcattaatatttgcoco-
caaatccaatgaagcctctgttgaactgttttaggcattagaagtattat-
caagtttctctgcaagcttgcattttccactccacctttagtcttttccacaacatct-
taccocctcccaaattcgacttttgcctatattgttagccaagttccaggaccattgtct-
catggttctgtcacgctaaattcatcatctgacgtgagaatcgctccaatattaaattcaat-
tactattcaaatccacagacctgtctaattgtttagcggcatgaagaagcttgggact-
tattaaaggacaaaggcattagaaccatataaagctcgagaatgctgggaattgacgggtt-
caattattgggagttaccttgcagagcaaccaacagatgaatccttcgaaa-
catttgcagataatgtagcttcatactggcattaccacgggtggaagccttgttggga-
aagtgcttgatgacagtttccgtttagtgggatcaaaagcattacgcgttgttgatgcctc-
cacttccctacgaaaccaaacagccatcctcagggtctctatctgatgttaggaagg-
tatgtgggccttcaatccctgcagaagaaggatcgaatgagtgaggcggcccatgcgaattc

bioRxiv preprint doi: <https://doi.org/10.1101/000000>; this version posted October 1, 2014. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Figure 5: Amino acid sequence of the hybrid protein PamHNL5xGOX, derived from the nucleic acid sequence (figure 4).

MQTLLVSSLVSLAAALPHYIRSNGIEAYNATDTSSEGSYDYVIGGGTSGCPLAATL
SEKYKVLLEGGTATIEYPNTLTADGFAYNLQQDDGKTPVERFVSEDGIDNVRARI
LGGTTIINAGVYARANISFYSGTGEWDLVLNKTIEWVEDAIVKPNNQSWQSVIG
EGFLEAGILPDNGFSLDHEAGTRLTGSTFDNNGTRHAADELLNKGDPNNLLVAVQA
SVEKILFSSNTSNLSAIGVIYTDSDGNSHQAFVRGNGEVIVSAGTIGTPQLLLSGVG
PESYLSSLNITVVQPNPYVGQFVYDNPRNFILPPNPIEASVVTVLGIRSDYYQVSL
SLPFSTPPFSLFPTTSYPLPNSTFAHIVSQVPGPLSHGVSITLNSSSDVRIAPNIKFN
YSNSTDLANCVSGMKKLGDLLRTKALEPYKARDVLGIDGFNYLGVPLPENQTDAS
FETFCLDNVASYWHYHGGS LVGKVLDDSFVRMGIKALRVVDASTFFPYEPNSHPQG
FYLMLGRYVGLQILQERSMO

2024.10.23 09:02:01

Figure 6: Comparison of the amino acid sequences of *Prunus amygdalus* HNL5 and of the hybrid protein PamHNL5xGOX. Sequence parts of *Aspergillus niger* glucose oxidase are underlined. Sequence regions having no significant homology between the two proteins are printed in italics, and the signal peptides are printed in bold type.

```

PamHNL5Gox 1  mqtlvsslv vsiaalphy lrngieg-----YNATDTSS
PamHNL5      1  mekstmsvll fvthlvhl qysevhsian tsahdsfyik fVYNATDTSS

PamHNL5Gox 37  EGSYDYIVIG GGTSGCPLAA TLSEKYKvLL LERGtIATEY PNTLTADGFA
PamHNL5      51  EGSYDYIVIG GGTSGCPLAA TLSEKYKvLL LERGtIATEY PNTLTADGFA

PamHNL5Gox 87  YNLQQQDDGK TPVERFVSED GIDNVRARIL GGTTIINAGV YARANISFYs
PamHNL5      101  YNLQQQDDGK TPVERFVSED GIDNVRARIL GGTTIINAGV YARANISFYs

PamHNL5Gox 137  QTGIEWDLDL VNkTYEWVED AIVVKPNnQS WQSVIGEGFL EAGILPDNGF
PamHNL5      151  QTGIEWDLDL VNkTYEWVED AIVVKPNnQS WQSVIGEGFL EAGILPDNGF

PamHNL5Gox 187  SLDHEAGTRL TGSTFDNNGT RHAADeLLNK GdPNLLVAV QASVEKILFS
PamHNL5      201  SLDHEAGTRL TGSTFDNNGT RHAADeLLNK GdPNLLVAV QASVEKILFS

PamHNL5Gox 237  SNTsNLsAIG VIYTDSDGNS HQAFVRGNGE VIVSAGTIGT PQLLLLSGVG
PamHNL5      251  SNTsNLsAIG VIYTDSDGNS HQAFVRGNGE VIVSAGTIGT PQLLLLSGVG

PamHNL5Gox 287  PESYLSSLNi TVVQPNPYVG QFVYDNPRNF INILPPNPIE ASVVTVLGIR
PamHNL5      301  PESYLSSLNi TVVQPNPYVG QFVYDNPRNF INILPPNPIE ASVVTVLGIR

PamHNL5Gox 337  SDYYQVSLSS LPFSTPPFSL FPtTSYPLPN STFAHIVSQV PGPLSHGSVT
PamHNL5      351  SDYYQVSLSS LPFSTPPFSL FPtTSYPLPN STFAHIVSQV PGPLSHGSVT

PamHNL5Gox 387  LNSSSDVRIA PNikFNYYsN STDLANCVSG MKKLGDLLRT KALEPYKARD
PamHNL5      401  LNSSSDVRIA PNikFNYYsN STDLANCVSG MKKLGDLLRT KALEPYKARD

PamHNL5Gox 437  VLGI DGfNYL GVPLPENQTD DASfETfCLD NVASyWHYHG GsLVGKVLDd
PamHNL5      451  VLGI DGfNYL GVPLPENQTD DASfETfCLD NVASyWHYHG GsLVGKVLDd

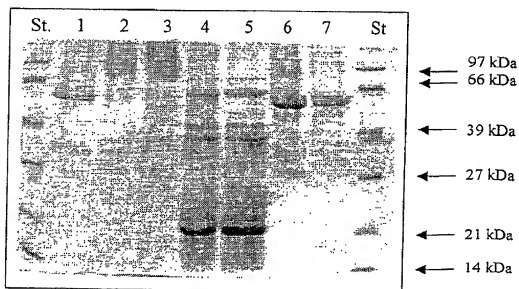
PamHNL5Gox 487  SFRVMGIKAL RVVDASTFPY EPNSHPQGfY LMLGRYVGLQ ILQERSmg-
PamHNL5      501  SFRVMGIKAL RVVDASTFPY EPNSHPQGfY LMLGRYVGLQ ILQERSrle

PamHNL5Gox 535 -----
PamHNL5      551 aihnlqesm

```

1004523.011602

Figure 7: Analysis of HNL preparations by SDS PAGE.
Details are described in example 11.



10046232.011602

Figure 8:

Nucleotide sequence of the *Prunus amygdalus HNL1* gene obtained by PCR amplification.

ATGAGAAATCAACAATGTCAGCTATACTGTTGGTGTATACATTTTTGTCCTCC
ATCTTCAATATTCTGAGGTCCACT
CGCTTGCCACGACTTCTGATCATGgtaaatcacttcaaccgtaattcaaaacacccaaaagg-
caatcaaaaagaaaacg
gaaaaaagtgtaagaaaagcagatatagagcctgcatagatgcatgctatatacttttaaaaaactcttgcctt
gagatttgcagATTTTAGCTACCTGAGCTTTGCATACGACGCCACTGATCTA-
GAGTTGGAAGGATCATATGACTACGT
TATAGITGGCGGAGGAACATCAGGGTGTCCATTGGCAGCAACTTTATCAGAAAA
ATACAAGGTGCTCGTTCTCGAAAGG
GGCAGTCTTCCGACAGCATATCCCAACGTCTTGACTGCAGATGGGTTTGATAT
AATCTCCAGCAAGAAGATGATGGAA
AGACACCGGTCGAAAGGTTCTGTGTCGGAAGATGGTATTGATAATGTACGGGGC
AGGGTGCTCGGTGGCACAAGCATTAT
CAATGCCGGTGTCTACGCCAGAGCTAACACCTCAATCTATAGTGCATCAGGAGT
TGATTGGGACATGGATTGGTTAAT
CAGACATATGAGTGGGTTGAAGACACTATTGTGTACAAGCCAAATTCTCAATCTT
GGCAGTCTGTTACAAAACTGCAT
TCTTGGAGGCTGGTGTTCATCCAACCATGGATTAGTTTAGATCATGAAGAA
GAACTAGAATTACCGGCTCAACTTT
TGACAACAAGGGAACGAGACATGCAGCTGATGAACCTTCTAATAAAGGAAACTC
TAACAACCTTGCAGTTGGAGTTCAT
GCCTCAGTAGAGAAGATCATCTTCTCCAATGCACCAAGgtatgtgcatgcatgcaatcaa-
aattaatattgtcattt
taaaacactagcaggagccaaggctcggagtagcaataaaattcattatttcttggattgttgaatgatta
taagctttctgaatgtagGTTTGACAGCTACAGGAGTCATATATAGGATTCTAATG-
GAACGCCCTCACCAAGCATT
GTACGCAGTAAGGGAGAAGTTATCGTGAGTGCAGGGACAATTGGGACCCCTCA
ACTTCTACTACTTAGCGGTGTTGGGC

104622-01602

CAGAGTCTTACCTATCATCTCTAAATATTCCAGTTGTTCTTTCCCATCCTTACGT
GGACAGTTTCTGCATGACAATCC
TCGTAATTTTCATTAACATTTTGCCCCAAATCCAATTGAACCCACAATGTAACTG
YTCTAGGCATTTCAAACGATTTT
TACCAATGTTCTTTCTCGAGCTTGCCATTTAGAACTCCACCCTTCGGTTTTTTCC
CTAGTGCATCTTATCCCCTGCCAA
ATTGCACTTTTGCTCACTTTGCTAGCAAAGTGGCAGGACCTTTATCATATGGTTC
TCTCACACTGAAATCATCCTCCAA
TGTGAGAGTCAGTCCAAATGTCAAATTTAATTACTATTCAAATCTGACAGATCTTT
CTCATTGTGTTAGCGGCATGAAG
AAGATTGGTGAACCTTTGAGCACAGACGCATTAACCATATAAAGTTGAAGATT
TGCCGGGTGTAGAAGGTTTTAATA
TTTTGGGAATCCCTTTGCCAAAGGACCAACAGATGATGCAGCCTTCGAAACAT
TTTGCCGAGAATCAGTAGCCTCATA
TTGGCACTACCACGGTGGATGCCCTTGTTGAAAGGTGCTTGATGGTGATTTCCG
TGTTACAGGGATCAATGCATTACGC
GTTGTTGATGGCTCAACATTCCCTTACACACCAGCGAGCCACCCTCAGGGCTTC
TATCTGATGTTAGGGAAGtatgta
caaatctcaataattatgttgagtggtctgttgtaataaactctatgccatattcttctcatccttcca
tttttgccatgggcagGTATGTGGGCATTAAATTCTGCAAGAAAGATCAGCTTCA-
GATCTAAAAATCTTGGATTCC
CTCAAGTCAGCAGCATCTTGCTCTTAACT

Figure 9:

Amino acid sequence of *Prunus amygdalus* hydroxynitrile lyase (HNL1), derived from the nucleotide sequence of the HNL1 gene.

MEKSTMSAILLVLYIFVLHLQYSEVHSLATTSDDHDFSYSLSFAYDATDLELEGSY
DYVIVGGGTSGCPLAATLSEKYKVLVLERGSLPTAYPNVLTADGFVYNLQQE
DDGKTPVERFVSEDGIDNVRGRVLGGTSHNAGVYARANTSIYSASGVDWDM
DLVNQTYEWVEDTIVYKPNSSQSWQSVTKTAFLLEAGVHPNHGFSLDHEEGTRI
TGSTFDNKGTRHAADELLNKGNSNNI.RVGVHASVEKIIIFSNAPGLTATGVIYR
DSNGTPHQAFVRSKGEVIVSAGTIGTPQLLLLSGVGPESYLSLNPVVLSHPY
VGQFLHDNPRNFINILPPNPIEPTIVTLGISNDFYQCSFSSLPFTTPPFGFFPS
ASTPLPNSTFAHFASKVAGPLSYGSLTLKSSSNRVSPNVKFNYYSNLTDLSHC
VSGMKKIGELLSTDALKPYKVEDLPGVEGFNILGIPLPKDQTDAAAFETFCR
ESVASYWHYHGGCLVGKVLGDGDFRVGTGINALRVVDGSTFPYTPASHPQGFYL
MLGRYVGKILQERSASDLKILDSLKSAASLV

10046232.014602